Preliminary evaluation of a two-interval, two-alternative infant behavioral testing procedure[a)]

Jenna Browning
Department of Allied Health Sciences, The University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599
jenna_browning@med.unc.edu

Emily Buss
Department of Otolaryngology/Head and Neck Surgery, The University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599
ebuss@med.unc.edu

Lori J. Leibold[b)]
Department of Allied Health Sciences, The University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599
leibold@med.unc.edu

Abstract: The purpose of this study was to evaluate the feasibility and efficiency of an observer-based, two-interval forced-choice infant psychophysical testing procedure. Ten of 11 infants (7–9 months of age) achieved a criterion of 80%-correct detection of a 50-dB sound pressure level noise band in a single testing session. Fewer trials were needed to reach criterion using the two-interval procedure than previously reported for the single-interval observer-based psychophysical procedure [Olsho, Koch, Halpin, and Carter (**1987**). Devel. Psychol. **23**, 627–640]. These results provide preliminary evidence that the two-interval procedure is feasible and efficient while controlling for observer and listener response bias.

PACS numbers: 43.66.Yw, 43.66.Sr, 43.66.Gf [QJF]
Date Received: June 5, 2014 **Date Accepted:** August 5, 2014

1. Introduction

Current behavioral methods used to test infants in the clinic or laboratory can be influenced by response bias of the listener and/or the examiner. Visual Reinforcement Audiometry (VRA) provides some control for examiner bias in the audiology clinic when control trials are used (e.g., Widen *et al.*, 2000), but does not account for listener bias. Moreover, conditioned head-turn procedures such as VRA are generally not effective for typically developing infants younger than about 5 months (e.g., Moore *et al.*, 1977), or for children with severe developmental disabilities (e.g., Gans and Gans, 1993), because these listeners do not reliably produce clear, short-latency head turns in response to sound. Olsho *et al.* (1987) developed a single-interval, observer-based procedure for testing infants in the laboratory. This is an operant-conditioning procedure in which the examiner does not know whether or not a signal is presented on a given trial. Signal-present trials are identified based on observations of the infant's behavior, so the examiner in this paradigm is referred to as the observer. One advantage of the observer-based procedure over conditioned head-turn procedures such as VRA is that the observer's decision regarding whether or not a signal was presented

[a)]These results were presented at the 41st Annual Scientific and Technology Conference in Scottsdale, AZ in March, 2014.
[b)]Author to whom correspondence should be addressed.

can be based on a wide variety of infant responses to sound (e.g., eye movements or a change in motor activity). The observer is blind to trial type (i.e., signal or no-signal); if the observer can reliably judge whether or not a sound has been presented based only on the infant's behavior, the underlying assumption is that the infant must have heard the sound. This method has been successfully used to test infants as young as 1 month of age (e.g., Werner and Gillenwater, 1990).

Observer bias is evaluated in the single-interval observer-based procedure, but listener bias remains a potential issue that can limit comparisons in adaptive threshold estimates between infants and older listeners. Specifically, adults are typically conservative in their response bias during single-interval procedures in which the temporal interval is undefined (e.g., Marshall and Jesteadt, 1986); this is true whether the listener responds directly (e.g., by pressing a button) or the observer-based method is used to identify signal-present trials (e.g., Leibold and Werner, 2006). In contrast, observer/infant teams tend to be unbiased in the single-interval observer-based procedure (e.g., Leibold and Werner, 2006). These age effects in response bias can result in differences in d' at threshold between infants and adults (e.g., Leibold and Werner, 2006).

The purpose of the present study was to evaluate the feasibility and efficiency of a two-interval forced-choice (2IFC) adaptation of the observer-based method to address the problem of response bias. Feasibility was assessed by determining the proportion of infants who successfully performed the task, reaching a criterion of 80%-correct detection performance in a single session. Efficiency was examined by comparing the number of trials required to reach criterion performance with the two-interval task to the number previously reported with the single-interval procedure (e.g., Werner and Boike, 2001).

2. Method

2.1 Listeners

Eleven 7- to 9-month-old infants participated in this study. Inclusion criteria included: (1) No risk factors for hearing loss as assessed by parent report, (2) fewer than three episodes of otitis media, (3) not receiving treatment for otitis media in the past two weeks, and (4) healthy on the day of testing. In addition, all 11 infants passed screening tympanometry in both ears immediately following testing. Peak admittance of at least 0.2 mmhos at a pressure between −200 and 50 daPa was required to pass the screening (e.g., Roush *et al.*, 1995).

2.2 Stimuli and equipment

Based on Werner and Boike (2001), the stimulus was a broadband noise, low-pass filtered at 6 kHz (48 dB/octave). Stimulus duration was 500 ms, with 10-ms onset and offset ramps. Custom software (MATLAB, MathWorks, Natick, MA) was used to generate and control the presentation of the stimulus. The stimulus was amplified (Applied Research and Technology, SLA4, Niagara Falls, NY) and presented using a loudspeaker (Monitor Audio, Monitor 4, Pickering, Canada). During testing, the infant sat on a parent's lap, positioned 3.3 ft from the loudspeaker in the sound field of a 7 × 7 ft, double-walled, sound-treated booth at approximately 0° azimuth and 0° elevation.

2.3 Procedure

Infants were tested using an observer-based, 2IFC psychophysical procedure. This is an operant-conditioning procedure, based on the observer-based, single-interval procedure developed by Werner and colleagues (e.g., Olsho *et al.*, 1987; Werner, 1995). An assistant sat to the right of the infant/parent in the sound booth and manipulated quiet toys in order to keep the infant facing toward midline. In order to prevent the assistant and the parent in the sound booth from hearing the noise stimulus and influencing the infant's response, the two adults in the booth wore noise-isolating earphones (Etymōtic mc5, Elk Grove Village, IL) that delivered masking sounds as well as noise-reduction earmuffs (Bilsom Thunder T3, Smithfield, RI). To the infant's left (90° azimuth) were two mechanical toys with lights inside dark Plexiglas boxes. An observer sat outside the booth and initiated trials when the infant was quiet and facing midline.

The signal was presented on each trial during one of two 500-ms temporal intervals, separated by a 300-ms inter-stimulus interval. Each interval was visually marked in real-time on a computer monitor, but the observer was not told which interval contained the signal. Following each trial, the observer was required to select the listening interval in which the signal occurred, based on the infant's behavior. The observer was given feedback after every trial indicating the signal interval.

The testing session lasted approximately 45 min. Sessions included a brief conditioning phase followed by a criterion phase. For both phases, the signal was presented at a supra-threshold level of 50 dB sound pressure level (SPL). The goal of the conditioning phase was to establish the relationship between the presentation of the stimulus and the mechanical toy reinforcement. The signal was always presented during the second interval during the conditioning phase. The infant was reinforced by the activation and illumination of a mechanical toy immediately after the end of each trial, regardless of the listener's response. This was done to facilitate response acquisition (e.g., Lancioni, 1980). The conditioning phase ended after a minimum of five trials, provided the observer judged that the relationship between the stimulus and reinforcer was established. Consistent with previous studies using the single-interval procedure (reviewed by Werner, 1995), the number of conditioning trials ranged from 5–9 for infants (mean = 6.2 trials).

During the criterion phase, the probability of the signal being presented in either interval was 0.50. The observer was prompted to enter a response after the second temporal interval. Reinforcement was provided to the infant only if the observer correctly identified the interval that contained the signal. For trials in which the signal was presented in the first interval, the toy reinforcer was activated immediately after the observer entered a response. For trials in which the signal was presented in the second interval, reinforcement was delayed by 800 ms following the observer's response. The rationale for inserting this delay was to maintain a consistent temporal relationship between the signal and reinforcer, regardless of which listening interval contained the signal. The criterion phase was completed when the observer correctly identified the interval that contained the signal on eight out of the previous ten trials.

3. Results and discussion

The primary goal of this study was to evaluate the feasibility and efficiency of an observer-based, 2IFC observer-based infant psychophysical procedure. The present findings support feasibility. Ten of 11 infants were able to perform the task, reaching the criterion of 80% correct in a single testing session. The two-interval procedure was also efficient. Infants tested with the two-interval procedure required fewer trials to reach criterion performance than previously reported for infants around the same age tested with the single-interval observer-based procedure (e.g., Leibold and Werner, 2006; Werner and Boike, 2001). Figure 1 shows the number of trials required to reach criterion performance for individual infants tested in the present study (open circles). On average, infants required 11 trials to complete the criterion phase with the two-interval procedure (solid horizontal line). In contrast, 6- to 9-month-old infants tested by Werner and Boike (2001) and 7- to 9-month-olds tested by Leibold and Werner (2006) with the single-interval procedure needed 20 and 21 trials, respectively, to reach an 80%-correct criterion.

The motivation for evaluating an observer-based, two-interval procedure was to reduce the influence of response bias on adaptive threshold estimates while retaining the advantages inherent in the observer-based technique. The effect of response bias on thresholds estimated adaptively using single-interval procedures has been a concern in the fields of developmental psychoacoustics (e.g., Olsho *et al.*, 1987; Leibold and Werner, 2006) and clinical audiology (e.g., Marshall and Jesteadt, 1986). For example, Leibold and Werner (2006) examined infant–adult differences in detection thresholds

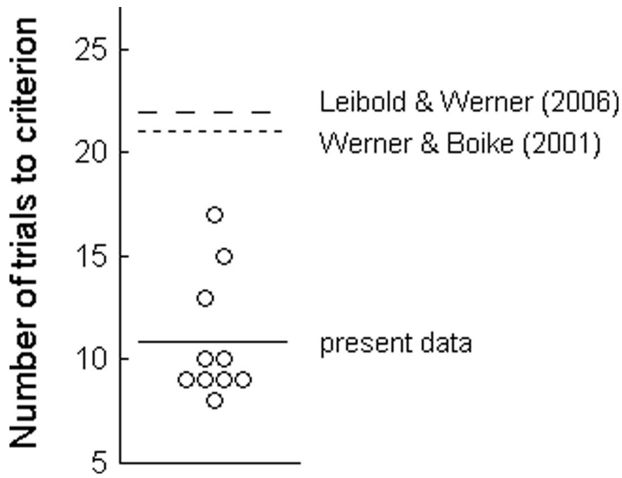

Fig. 1. The number of trials to reach criterion to achieve 80%-correct detection of a 50-dB SPL noise band presented in quiet is plotted for individual infants ($n = 10$). The solid horizontal line shows the average number of trials needed for infants to reach criterion in the present study. The large and small dashed horizontal lines show the average number of criterion trials for infants tested by Leibold and Werner (2006) and Werner and Boike (2001), respectively.

estimated adaptively with the single-interval observer-based procedure. An analysis of d' near threshold indicated that adults were operating at a higher sensitivity than infants, suggesting that the adaptive thresholds underestimated the true difference in sensitivity between infants and adults. A primary benefit of the two-interval paradigm is that it is believed to reduce or eliminate response bias (e.g., Green and Swets, 1973), thus permitting direct comparison across different ages and/or stimuli using efficient adaptive procedures. Future studies are planned to compare adaptive threshold estimates measured with the two- and single-interval procedures. Thresholds for infants are expected to be similar across the procedures, consistent with previous findings of no response bias (e.g., Leibold and Werner, 2006). Thresholds for adults are expected to be lower in the two-interval than in the single-interval procedure due to the elimination of response bias in the two-interval procedure.

The two-interval procedure offers other potential advantages over the single-interval procedure. Most notable is the smaller number of trials needed to reach criterion performance with the two-interval than with the single-interval procedure. This is particularly beneficial because infants habituate and stop responding to reinforcement over time (e.g., reviewed by Werner, 1995). Another potential benefit of the two-interval task is increased ability of the observer to identify behaviors associated with signal detection. In the single-interval task the observer must determine the behavior associated with signal detection for each listener, based on task feedback over trials. This skill is developed with extensive observer training and experience, and it works best if the listener is very consistent in his/her response. With the two-interval adaptation, the observer can base each response on a *change* in the listener's behavior during the trial, irrespective of the nature of that response. This reduces the memory load on the experimenter, allowing a comparison of behaviors over a shorter time interval, and accommodates changes in the listener's behavior over the course of the experiment.

The two-interval adaptation of the observer-based procedure also represents an important innovation over previous two-interval infant testing procedures that have been used to collect infant behavioral data in the laboratory or audiology clinic (e.g., Schneider *et al.*, 1980). Previous two-alternative infant test procedures require that the infant not only detect the stimulus of interest, but that they also localize its source and orient toward it; the present procedure does not rely on sound source localization, and therefore is not limited by effects of age and/or hearing loss on the binaural system.

The present findings may have clinical implications. Pediatric audiologists are acutely aware of the importance of working quickly during infant behavioral assessments. Consistent with other procedures that rely on operant conditioning, VRA testing is limited by the number of times reinforcement can be used before habituation occurs. For example, in a study describing VRA outcomes for over 3000 infants, Widen *et al.* (2000) reported that the most common reason infants failed to complete the VRA protocol in a single testing session was habituation. Thus, even a modest decrease in the number of reinforced stimulus trials provides an opportunity to increase the amount of behavioral data collected during clinical visits. The two-interval observer-based procedure might also prove useful for obtaining reliable behavioral data from "difficult-to-test" listeners, including toddlers and children with severe physical and/or developmental disabilities. Clinical VRA is a conditioned head-turn procedure, requiring a developmental age greater than 5–6 months (e.g., Widen *et al.*, 2000; Gans and Gans, 1993). The observer-based method has been shown to be effective with very young typically developing infants (e.g., Werner and Gillenwater, 1990); thus, it might also be useful in children with a developmental age younger than required for VRA testing. The implementation of a two-interval paradigm retains the benefits associated with observer-based methods, while protecting against response bias. Improving our ability to collect reliable behavioral data in toddlers and hard-to-test children would be invaluable for the clinical treatment of these listeners.

4. Conclusions

(1) Preliminary data indicate that the 2IFC procedure may be a feasible and efficient approach for testing infants. Fewer training trials were required to reach criterion than previously reported using single-interval procedures.
(2) This approach has the potential to improve our ability to compare performance across age while controlling for response bias and allowing flexibility in the listener response (e.g., head turns, or changes in motor activity).
(3) Future studies are planned to determine the reliability of adaptive thresholds estimated using this procedure, and to evaluate the utility of this procedure in listeners who are difficult-to-test using conventional VRA testing (e.g., toddlers).

Acknowledgments

This work was sponsored by the NIH (Grant No. R01 DC011038). The authors would like to thank the members of the Human Auditory Development Laboratory for their contributions, particularly Crystal Taylor and Steve Lockhart.

References and links

Gans, D., and Gans, K. D. (**1993**). "Development of a hearing test protocol for profoundly involved multi-handicapped children," Ear Hear. **14**, 128–140.
Green, D. M., and Swets, J. A. (**1973**). *Signal Detection Theory and Psychophysics* (Krieger Publishing, Huntington, NY).
Lancioni, G. E. (**1980**). "Infant operant conditioning and its implication for early intervention," Psychol. Bull. **88**, 516–534.
Leibold, L. J., and Werner, L. A. (**2006**). "Effect of masker-frequency variability on the detection performance of infants and adults," J. Acoust. Soc. Am. **119**, 3960–3970.
Marshall, L., and Jesteadt, W. (**1986**). "Comparison of pure-tone audibility thresholds obtained with audiological and two-interval forced-choice procedures," J. Speech Hear. Res. **29**, 82–91.
Moore, J. M., Wilson, W. R., and Thompson, G. (**1977**). "Visual reinforcement of head-turn responses in infants under 12 months of age," J. Speech Hear. Dis. **42**, 328–334.
Olsho, L. W., Koch, E. G., Halpin, C. F., and Carter, E. A. (**1987**). "An observer-based psychoacoustic procedure for use with young infants," Dev. Psychol. **23**, 627–640.
Roush, J., Bryant, K., Mundy, M., Zeisel, S., and Roberts, J. (**1995**) "Developmental changes in static admittance and tympanometric width in infants and toddlers," J. Am. Aud. Audiol. **6**, 334–338.

Schneider, B. A., Trehub, S. E., and Bull, D. (**1980**). "High-frequency sensitivity in infants," Science **207**, 1003–1004.
Werner, L. A. (**1995**). "Observer-based approaches to human infant psychoacoustics," in *Methods in Comparative Psychoacoustics* (Birkhäuser, Basel), pp. 135–146.
Werner, L. A., and Boike, K. (**2001**). "Infants' sensitivity to broadband noise," J. Acoust. Soc. Am. **109**, 2103–2111.
Werner, L. A., and Gillenwater, J. M. (**1990**). "Pure-tone sensitivity of 2- to 5-week-old infants," Infant Behav. Dev. **13**, 355–375.
Widen, J. E., Folsom, R. C., Cone-Wesson, B., Carty, L., Dunnell, J. J., Koebsell, K., Levi, A., Mancl, L., Ohlrich, B., Trouba, S., Gorga, M. P., Sininger, Y. S., Vohr, B. R., and Norton, S. J. (**2000**). "Identification of neonatal hearing impairment: hearing status at 8 to 12 months corrected age using a visual reinforcement audiometry protocol," Ear Hear. **21**, 471–487.